PNAS

www.pnas.org

Supplementary Information for

Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle

Parijat Majumder[a], Till Rudack[b], Florian Beck[a], Radostin Danev[a,1], Günter Pfeifer[a], István Nagy[a], and Wolfgang Baumeister[a,2]

[a]Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany;
[b]Department of Biophysics, Ruhr University Bochum, 44801 Bochum, Germany;
[1]Present address: Graduate School of Medicine, University of Tokyo, 113-0033 Tokyo, Japan

[2]To whom correspondence should be addressed. Email: baumeist@biochem.mpg.de

This PDF file includes:

    Supplementary text
    Figs. S1 to S6
    Tables S1 to S2
    Captions for movies S1 to S2
    References for SI reference citations

Other supplementary materials for this manuscript include the following:

    Movies S1 to S2

SI Materials and Methods

Expression and purification of *A. fulgidus* PAN. A codon optimized version of the *A. fulgidus pan* gene was cloned into a modified pCDF-Duet vector containing a T7 promoter, an N-terminal 8xHis tag followed by a solubilisation tag Ta0895 and a Tobacco Etch Virus (TEV)-protease cleavage site between the Ta0895 tag and the *pan* gene. Protein was expressed in *E. coli* BL21-CodonPlus (DE3)-RIL cells. The expression cells were grown to an optical density at 600 nm (OD_{600}) of 0.6-0.8, and induced with 1 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) for 20 hours at 18°C. Cells were harvested by centrifugation at 6,000 rpm for 10 min, flash frozen in liquid nitrogen, and stored at -20°C for at least 2 hours. PAN was purified using a protocol modified from ref.(1). Frozen cells were resuspended in buffer A (50 mM Tris, pH 8.0, 150 mM KCl, 150 mM NaCl, 10% glycerol, 2 mM DTT, 1 mM EDTA and protease inhibitors) supplemented with 2 mg/ml lysozyme and benzonase, and lysed by sonication on ice (4 x 10 pulses with 2 min gaps, 60% duty cycle). The lysate was clarified by centrifugation at 20,000 rpm at 4°C for 30 min, and the cleared lysate was purified in batch using cOmplete His-tag purification resin (Roche). The PAN protein was eluted with 250 mM Imidazole in buffer A and was buffer exchanged using PD10 columns to buffer B (50 mM Tris, pH 8.0, 50 mM NaCl, 5% glycerol and 2 mM DTT). Contaminating proteins were removed by MonoQ anion exchange chromatography, using a linear salt gradient from 50mM to 1M NaCl in buffer B, and the PAN containing fractions were incubated with TEV-protease overnight on ice. For PAN hexamer assembly, the samples were incubated at 80°C for 15 min, and cooled on ice for 30 min. The TEV-protease aggregated upon heating and was removed by centrifugation at 10,000 rpm for 10 min, followed by gel filtration through a superose 6 column (GE Healthcare) equilibrated with buffer C (50 mM Tris, pH 8.0, 200 mM NaCl, 5% glycerol, and 1 mM DTT) (Fig. S1*A* and *B*).

Heterologous expression and purification of *A. fulgidus* CP. The *A. fulgidus psmA* gene, that encodes the proteasome α-subunit, was cloned into a modified pCDF-Duet vector containing a T7 promoter, and no affinity tag. An N-terminal truncated version of the *A. fulgidus psmB* gene (amino acids 11-213) was cloned into a modified pET-Duet vector containing a T7 promoter, and a C-terminal 6xHis tag. CP expression strains were generated by co-transforming the pCDF-Duet and pET-Duet plasmids into *E.coli* BL21-CodonPlus(DE3)-RIL cells. The transformed cells were grown to an OD_{600} of 0.8-0.9 and induced with 1 mM IPTG at 25°C overnight. Cells were harvested and lysed in 100mM Hepes, pH 7.5, 10mM DTT, and 1mM EDTA supplemented with 2 mg/ml lysozyme and benzonase. The cleared lysate was incubated at 85°C for 30 min, and then cooled on ice for 15 min. Heat induced aggregates were removed by centrifugation at 14,000 rpm for 10 min at 4°C, and the clarified supernatant was affinity purified in batch. Assembled CP thus obtained were further cleaned by gel filtration through a Superose 6 column (GE Healthcare) equilibrated with glycerol free buffer C (Fig. S1*A* and *B*).

***In vitro* reconstitution of PAN-proteasomes.** Purified PAN and CP were mixed in a molar ratio of 3:1 in rec buffer (50 mM MOPS, pH 7.1, 300 mM KCl, 100 mM NaCl, 30 mM $MgSO_4$, and 5 mM DTT). Nucleotide or nucleotide analogs were added (2 mM ATP/

2 mM ADP/2 mM ADP, 8 mM NaF, 2 mM AlCl$_3$/ 2 mM ATPγS), and the reconstitution mix was incubated at 70°C for 15 min. After cooling to room temperature for 2 min, complex formation was checked by negative stain following an established protocol (Fig. S1*C*)(2).

Cryo-EM sample preparation. PAN-proteasomes were assembled in 5 mM ATPγS and incubated with amphipol A8-35 (1: 2.3 w/w) for 5 min. The excess amphipol was then removed by 2 successive passages through spin desalting columns, equilibrated in grid buffer (50 mM MOPS pH 7.1, 200 mM KCl, 20 mM MgSO$_4$, 20 mM K-Glutamate, 1 mM ATPγS).

For cryo-EM, 4µl of amphipol coated PAN-CP complexes (~2 mg/ml) in grid buffer or 4 µl CP (1.2 mg/ml) in glycerol-free buffer C was applied to glow discharged Lacey grids (Lacey carbon film, 200-mesh Cu grid), blotted with a Vitrobot Mark III (FEI Company) using 5 s blotting time with 95% humidity at 5°C, and plunge frozen in liquid ethane-propane mixture, cooled by liquid nitrogen.

Electron microscopy data acquisition. Initially negative stain datasets were acquired semi-automatically with a Tecnai (FEI) F20 transmission electron microscope operating at 200 kV. Micrographs were recorded using an Eagle 4K x 4K CCD camera at a nominal magnification of 62,000X and an object pixel size of 1.78 Å/pixel.

Cryo-EM data was acquired with an FEI Titan Krios transmission electron microscope using Latitide S software (Gatan). Movie frames were recorded at a nominal magnification of 18,000X using a K2 Summit direct electron detector camera (Gatan) operating in super-resolution mode. A total dose of ~30 electrons was distributed over 30 frames (1 electron per frame) and the calibrated physical pixel size and the super-resolution pixel size were 1.34 Å and 0.67 Å, respectively. Images were recorded in a defocus range of -1.4 to -2.4 µm for the control CP dataset and -1.7 to -2.7 µm for the PAN-proteasome dataset.

Image processing, classification and refinement. Cryo-EM micrograph frames were aligned using MotionCor2 (3) and the contrast transfer function (CTF) for the aligned frames was determined with CTFFIND4 (4). Micrographs with a defocus outside the range 0.8-3.0 µm, or a measured resolution worse than 6 Å were then discarded.

For the PAN-proteasome datasets (I and II) 2,452 and 3,636 micrographs remained. Particles were picked with Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) using negative stain 2D class averages as reference, and were processed in RELION 2.1 (5). 285,085 and 603,629 particles were picked from datasets I and II respectively using a 384 pixel box size, which were scaled down by a factor of 2. The particles were sorted by reference free classification in 2D and then by 3D classification using a 60 Å down-filtered *T. acidophilum* CP as reference (6). Ultimately a total of 208,833 particles remained from the 2 datasets combined, and they were refined in 3D with an imposed C2 symmetry.

To improve the alignment, *in silico* density subtraction (7) was performed to generate pseudo single capped particles (PSCs). PSCs were merged and refined without symmetry and the x and y shifts from this refinement were used to re-extract unbinned particles using a box size of 384 pixels. Re-extraction caused PSCs to revert back to original particles. Therefore, for further processing, the particles were masked around the CP and one PAN, and were treated as PSCs. 3D local refinement of these particles resolved to 4.43Å by gold standard FSC = 0.143 criterion.

In silico density subtraction was performed to leave the PAN OB ring, and the AAA ring (referred to as AAAob density) and the AAAob particles were refined with local angular searches, limiting the rot, tilt and psi search ranges to +/- 15°. The refined AAAob density was classified without rotational alignment into 12 classes. Among these 12 classes, 5 classes showed rotated spiral staircase conformations (corresponding to 60.2% of the dataset) and these classes were processed individually (Fig. S3).

Each class thus identified was cleaned as described in Fig. S4. The AAAob particles were reverted back to original particles and locally refined as PSCs. The OB ring and coiled-coils (together referred to as OB-cc) was then classified by focused classification (without rotational alignment) using a spherical mask. The best resolved class was selected and the PSCs were again locally refined. To further resolve structural details within these classes, the particles were density subtracted into 4 parts - AAAob, AAA ring, 20Sαβ and the interface (comprising of CP α ring and PAN AAA ring), and each model was refined individually.

In order to generate a merged AAA-ring structure, AAA-ring reconstructions from the different classes were aligned to the best resolved class (state 2). The boxed particles were then rotated according to the corresponding aligned class volumes and further classified in 3D.The best resolved class containing 82,207 particles was finally refined in 3D with local angular searches.

For the *A. fulgidus* control CP dataset, 142,031 particles were picked from 425 micrographs using a box-size of 160. After several rounds of 2D and 3D classification, 55,235 particles remained. The particles were *in silico* density subtracted into half (20Sαβ), and upper and lower halves were then combined. After further classification in 3D, the remaining 105,384 particles of 20Sαβ were locally refined in 3D with imposed c7 symmetry.

Model building. A pseudo-atomic model of *A. fulgidus* PAN-proteasome was generated following the strategy employed earlier to model the 26S proteasome (8-10). Based on the crystal structure of the PAN N-domain [PDB: 2WG5, (11)] and the ATPase domain [PDB: 3H4M, (12)], we generated a homology model of PAN employing Modeler (13), and created the missing unresolved segments by *ab initio* structure prediction in Rosetta (14). Then the generated structural model of PAN, together with the crystal structure of the *A.fulgidus* 20Sαβ [PDB: 1J2Q, (15)] was fitted by rigid body docking into the density of the highest resolved state (state 2), and refined to the density using molecular dynamics flexible fitting (MDFF)(16). The MDFF simulations were prepared using

QwikMD (17), analyzed with VMD (18), and carried out with NAMD (19). The obtained state 2 model of the PAN-proteasome was used to initiate further MDFF simulations to fit the densities for the other four states. We thus obtained pseudo-atomic structural models of five distinct states of the PAN-proteasome.

For refinement of the model to high-resolution densities (merged AAA-ring and 20Sαβ), we used an optimized workflow iteratively combining MDFF with Monte Carlo backbone and sidechain rotamer search algorithm (20).

Model based measurements. Pore-1 loop heights and nucleotide pocket widths were measured as described for the 26S proteasome (10). α5 (Pore-1), α6 (Pore-2), and α4 helices were defined in the UCSF Chimera package (21), and the Pore-1 loop height was measured as the distance between the N-terminal tip of α5 helix and the plane above CP α-ring (residue 171 of chains A-G). The distance between the mid-point of α6 helix and N-terminal tip of α4 helix was considered the nucleotide pocket size, and the distance between the N-terminal tips of α5 helices from two adjacent subunits was considered as the separation between Inter Pore-1 helix tips (Table S1). For measurement of angles, planes were defined passing through the middle of the AAA ring (residue 190 of chains H-M), the middle of the OB-ring (residue 77 of chains H-M), and above the CP α-ring (residue 171 of chains A-G). Inter-plane angles were then measured in Chimera. For measurement of offsets, centroids were defined for the CP α-ring, PAN AAA-ring, and PAN OB-ring. The centroids were projected on the same xy plane (plane parallel to the CP α-ring plane defined above), and Euclidean distances were measured between them (Table S2).
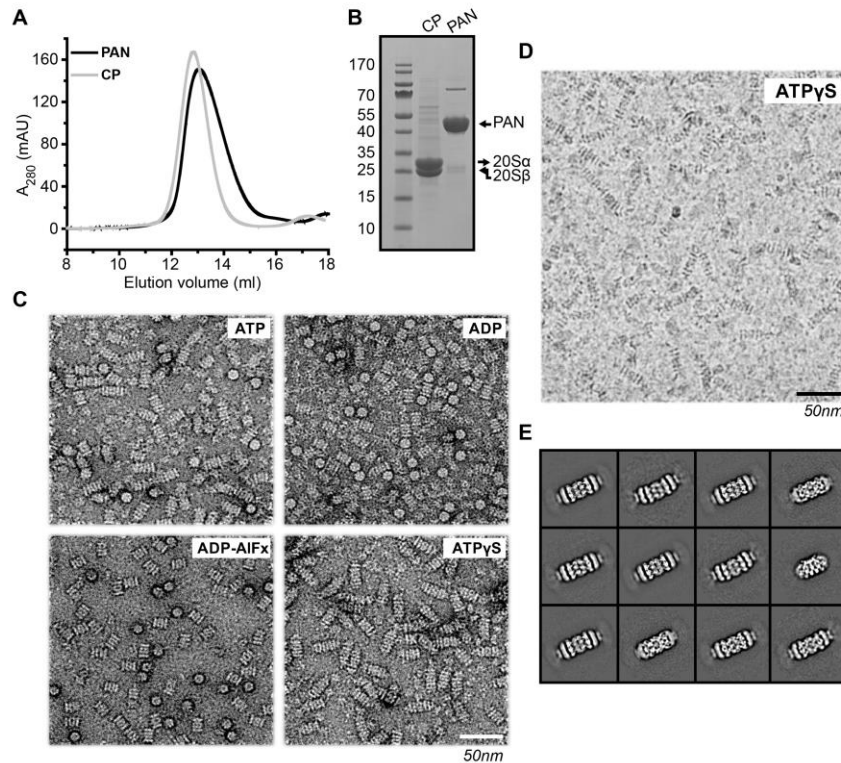
Figure S1. *In vitro* reconstitution and electron microscopy of *A. fulgidus* PAN-proteasomes. (*A*) Superose 6 gel filtration profiles of purified PAN (black) and CP (grey). (*B*) Coomassie stained SDS-PAGE showing the purified PAN and CP used for reconstitution. In case of the CP, two bands can be observed, corresponding to 20Sα and 20Sβ subunits. (*C*) PAN-CP interaction in presence of different nucleotides. PAN and CP were combined in a molar ratio of 3:1, and incubated in presence of either ATP, ADP, ADP-AlFx or ATPγS. Aliquots from different assembly reactions were assessed by negative stain. (*D*) A cryo-EM micrograph of ATPγS assembled PAN-proteasomes, down-filtered to 30Å. (*E*) Representative class averages obtained during reference-free 2D classification of the cryo-EM datasets.
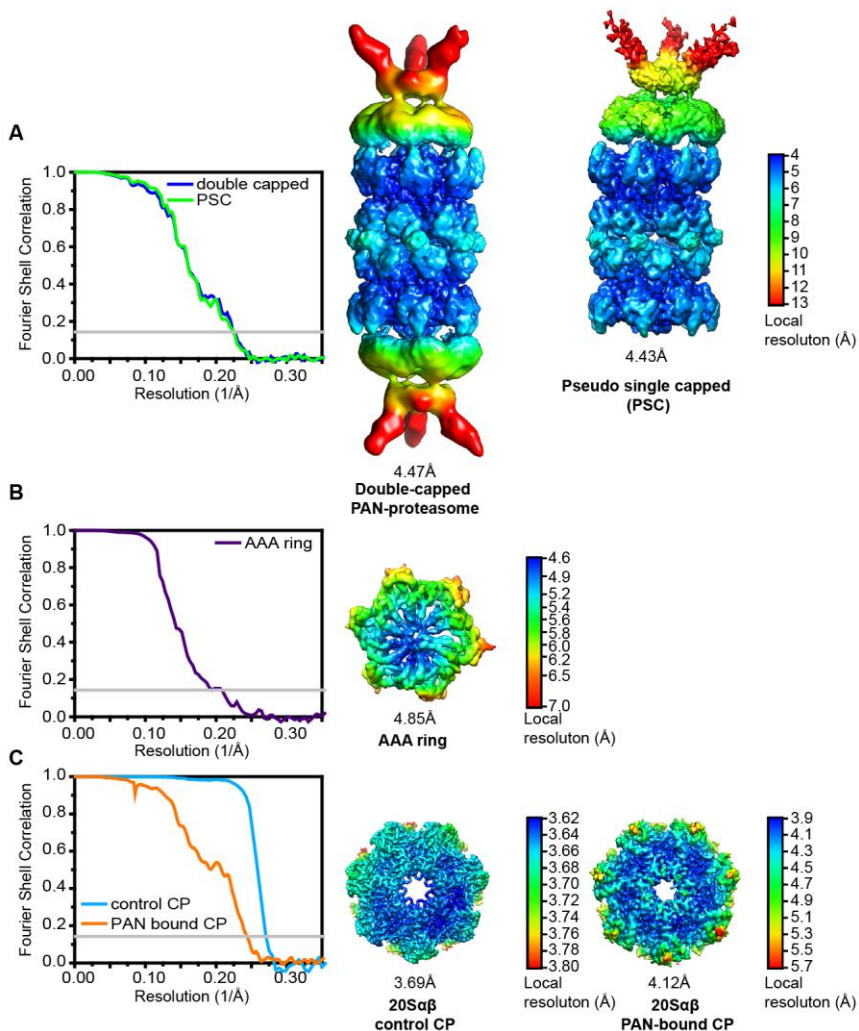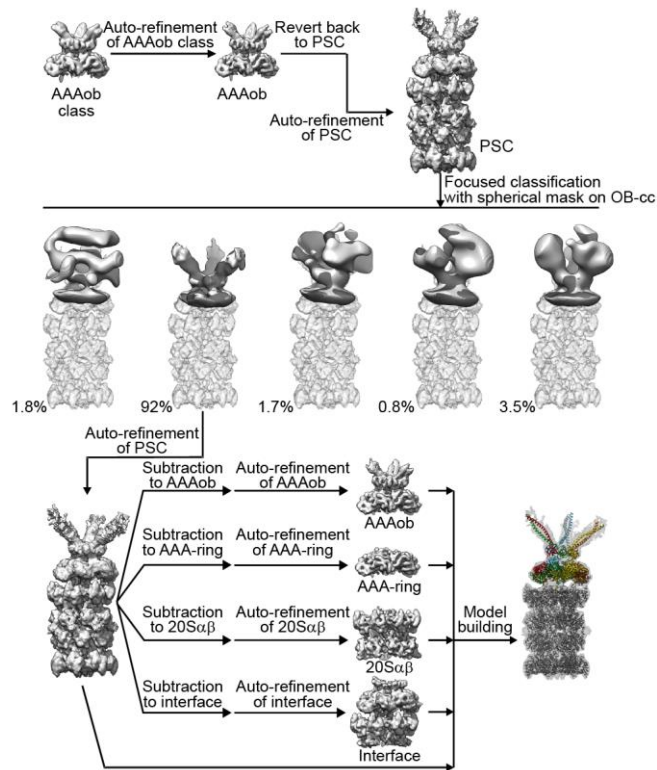
Figure S2. Overall and local resolution estimates of the PAN-proteasome. Fourier Shell Correlation (FSC) curves and density maps are displayed for (*A*) double capped and pseudo-single capped (PSC) PAN-proteasomes; (*B*) PAN AAA-ring (rotated and merged from all states); (*C*) *A. fulgidus* 20Sαβ control and from PAN-proteasomes (merged from all states). An FSC value of 0.143 is indicated by a grey line. *A-C* Values displayed below each density map show the corresponding overall resolution, estimated according to gold standard FSC = 0.143 criterion. The density maps are colored according to local resolution, as calculated by RELION 2.1, and color keys presented show the values of local resolution.
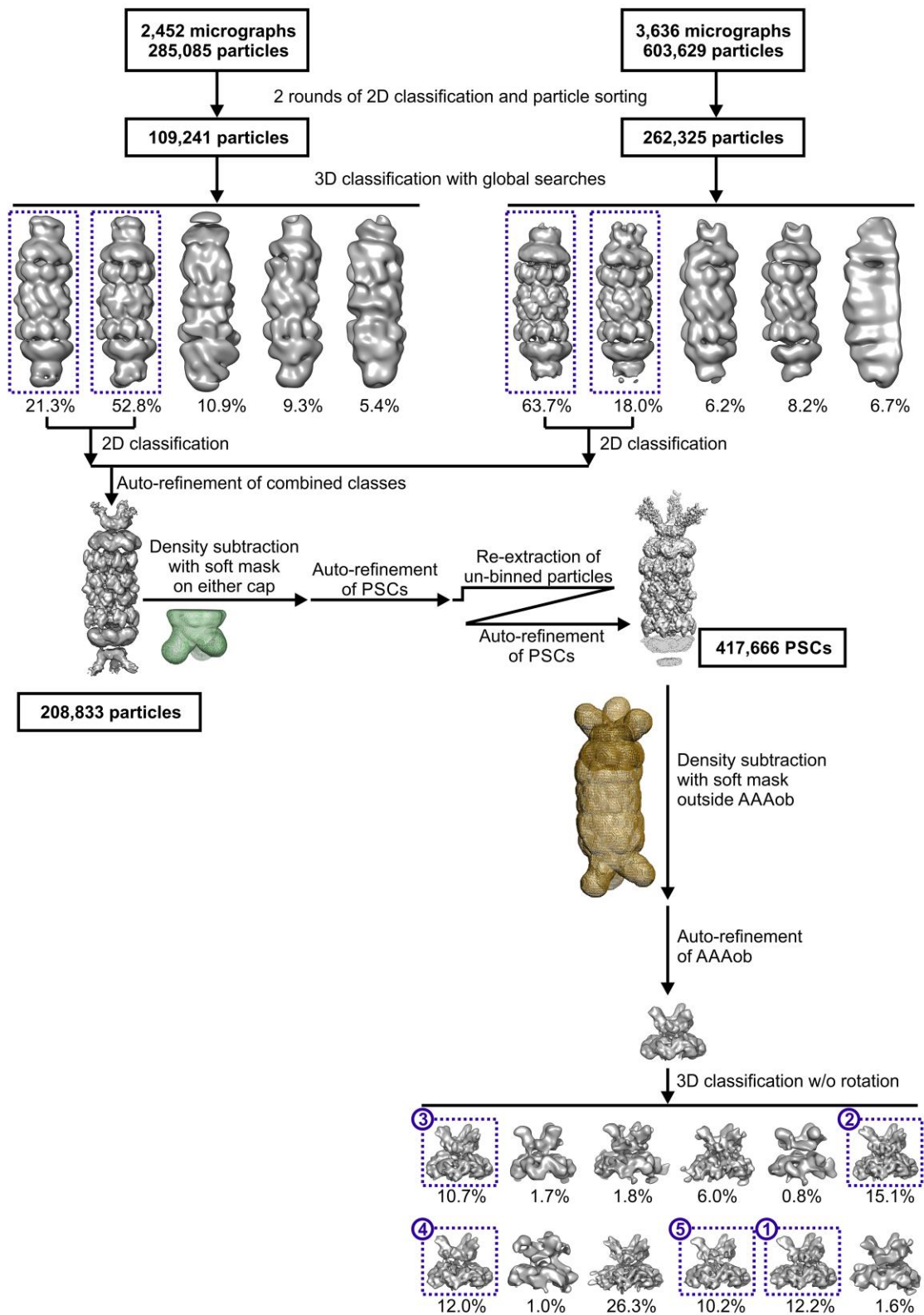
Figure S3. Classification of cryo-EM datasets. Image processing scheme followed to finally obtain distinct states of AAAob, which were refined and processed separately. The

selected classes are boxed in purple dashed lines and the numbers on the top left corner of each boxed AAAob class represent the state identification numbers.

**Figure S4.** Refinement and model generation. Scheme for *in silico* clean-up and refinement of individual classes. Each of the AAAob classes obtained (Fig. S3) are subjected to the same procedure, and for each class, 5 sets of densities are generated: PSC, AAAob, AAA-ring, 20Sαβ and interface. A model is built based on the PSC density and is refined using densities of AAAob, AAA-ring, 20Sαβ and the interface.
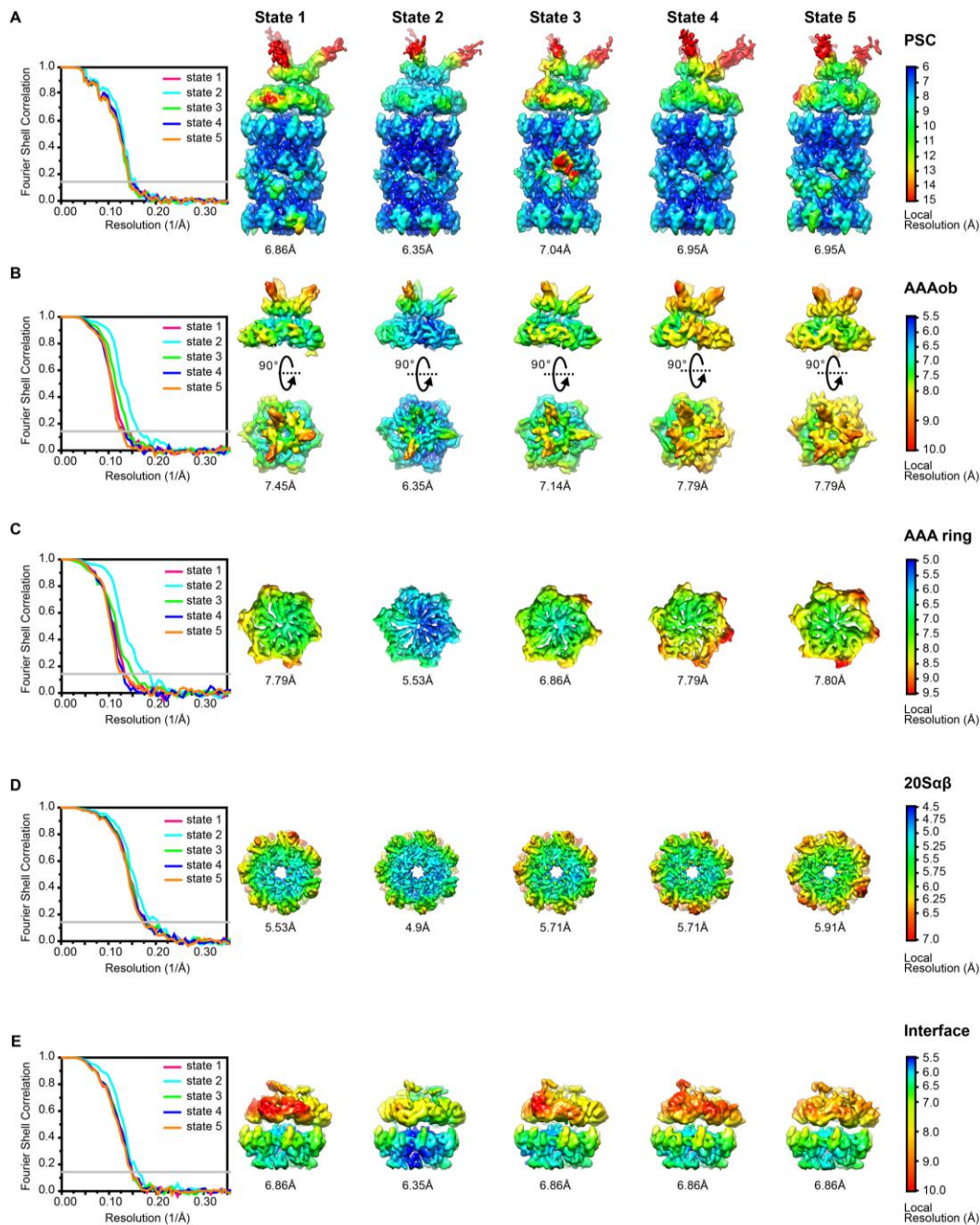
Figure S5. Overall and local resolution estimates of 5 states of the PAN-proteasome. For all 5 states, the FSC plots and density maps are displayed for (*A*) Pseudo-Single capped (PSC) particles, (*B*) PAN-AAAob, (*C*) PAN-AAA ring, (*D*) 20Sαβ and (*E*) PAN-CP interface or AAAα models. *A-E* The left panel shows an overlay of Fourier Shell Correlation (FSC) curves for the corrected maps from each of 5 states. An FSC value of 0.143 is indicated by a grey line. Density maps displayed from left to right correspond to states 1 through 5, and their overall resolutions (gold standard FSC = 0.143) are presented below. The density maps are colored according to local resolution, as

calculated by RELION 2.1, and the color keys presented on the right of each row show the values of local resolution for that particular row.
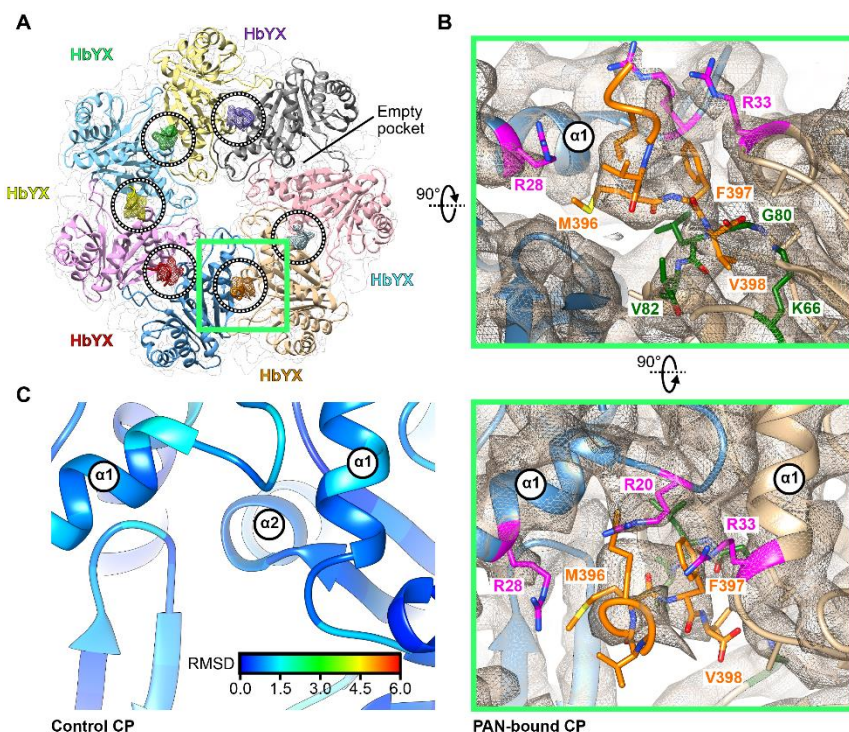
Figure S6. HbYX mediated interaction between PAN and CP. (*A*) Overview of six HbYX insertions in the CP inter-α-subunit pockets. Segmented HbYX densities in mesh form are shown with fitted models. HbYX densities and models are colored according to PAN protomer. In all five conformational states of the PAN-proteasome, the position of five HbYX insertions remain invariant. However, the density of the skyblue HbYX is the weakest and is incidentally the farthest from the PAN offset position. (*B*) Close-up of a representative HbYX binding pocket is shown from two different views. Atomic model of the CP is fitted into densities segmented from 20Sαβ. The HbYX model is colored orange. CP residues directly involved in HbYX interaction are colored dark green. (*C*) Model showing the HbYX binding pocket of a CP in absence of PAN (control CP). The model is generated by flexible fitting of the ~3.6Å control CP density map, and is colored according to the residue-wise Cα rmsd values. The rmsd values obtained by aligning the β ring of both control and PAN-bound CP models indicate that HbYX binding leads to only subtle changes in the inter-α-subunit pocket.

Table S1. Model based measurements elucidate the motion of PAN AAA-ring and OB-ring with respect to the CP.

Measurement	State 1	State 2	State 3	State 4	State 5
Offset between CP and AAA-ring (Å)	10.48	10.69	9.51	9.26	9.75
Offset between CP and OB-ring (Å)	17.46	15.17	11.20	9.13	12.58
Offset between AAA-ring and OB-ring (Å)	7.01	4.66	1.80	1.74	4.29
Angle between CP and AAA-ring (degrees)	1.38	3.00	4.46	4.58	3.00
Angle between CP and OB-ring (degrees)	9.10	12.84	14.44	13.05	6.44

Distances were measured based on the near-atomic models of the PAN-proteasome in different states. (Described in SI Materials and Methods)

Table S2. Model based measurements illustrate the geometry of the PAN AAA-staircase.

Measurement	PAN protomer	State 1	State 2	State 3	State 4	State 5
Pore-1 loop height (Å)	1	47.9	44.7	41.0	37.0	32.6
	2	56.7	49.6	46.4	42.2	35.6
	3	33.0	57.5	50.5	46.1	40.3
	4	36.1	32.5	58.2	50.3	45.3
	5	40.2	36.8	33.5	58.0	48.2
	6	44.5	40.4	36.7	33.7	55.7
Inter Pore-1 helix tip (Å)	1-2	11.2	9.9	9.5	9.8	10.4
	2-3	28.3	10.7	10.5	9.8	9.8
	3-4	10.8	28.8	10.4	10.6	10.4
	4-5	9.7	10.3	27.9	11.5	10.6
	5-6	10.2	9.6	10.7	27.6	10.4
	6-1	10.2	9.5	9.8	10.7	27.2
Nucleotide pocket (Å)	1-2	12.1	11.9	11.7	12.4	12.4
	2-3	25.7	11.3	12.2	11.8	12.3
	3-4	13.2	28.8	12.3	12.5	14.3
	4-5	14.1	13.7	28.9	13.2	12.9
	5-6	12.8	13.9	14.1	29.2	12.8
	6-1	12.1	12.0	12.6	13.8	26.5

Distances were measured based on the near-atomic models of the PAN-proteasome in different states. (Described in SI Materials and Methods)

Movie S1. Conformational states of the PAN-proteasome demonstrate an around-the-ring ATPase cycle. The 5 conformational states of the PAN-proteasome exhibit 5 distinct rotated spiral-staircase conformations of the AAA-ring. The AAA-ring movements are coupled with motions of the OB-ring and coiled-coils.

Movie S2. Conformational changes in the CP gate that occur upon PAN binding. PAN binding stabilizes the CP gate in an open conformation. The otherwise unstructured N-terminal tails of CP α-subunits form loops ('gate loops') that bring Y26 and D9 of one subunit close to P17 and Y8 of the adjacent subunit. In absence of PAN, these gate loops are not observed. However, inter-subunit Y26-P17 interactions are found lining the pore of the CP axial channel.

References

1.	Zwickl P, Ng D, Woo KM, Klenk HP, & Goldberg AL (1999) An archaebacterial ATPase, homologous to ATPases in the eukaryotic 26 S proteasome, activates protein breakdown by 20 S proteasomes. *The Journal of biological chemistry* 274(37):26008-26014.
2.	Booth DS, Avila-Sakar A, & Cheng Y (2011) Visualizing proteins and macromolecular complexes by negative stain EM: from grid preparation to image acquisition. *Journal of visualized experiments : JoVE* (58).
3.	Zheng SQ*, et al.* (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods* 14(4):331-332.
4.	Rohou A & Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of structural biology* 192(2):216-221.
5.	Scheres SH (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology* 180(3):519-530.
6.	Danev R, Tegunov D, & Baumeister W (2017) Using the Volta phase plate with defocus for cryo-EM single particle analysis. *eLife* 6.
7.	Bai XC, Rajendra E, Yang G, Shi Y, & Scheres SH (2015) Sampling the conformational space of the catalytic subunit of human gamma-secretase. *eLife* 4.
8.	Guo Q*, et al.* (2018) In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. *Cell* 172(4):696-705 e612.
9.	Schweitzer A*, et al.* (2016) Structure of the human 26S proteasome at a resolution of 3.9 A. *Proceedings of the National Academy of Sciences of the United States of America* 113(28):7816-7821.
10.	Wehmer M*, et al.* (2017) Structural insights into the functional cycle of the ATPase module of the 26S proteasome. *Proceedings of the National Academy of Sciences of the United States of America* 114(6):1305-1310.
11.	Djuranovic S*, et al.* (2009) Structure and activity of the N-terminal substrate recognition domains in proteasomal ATPases. *Molecular cell* 34(5):580-590.
12.	Zhang F*, et al.* (2009) Structural insights into the regulatory particle of the proteasome from Methanocaldococcus jannaschii. *Molecular cell* 34(4):473-484.

13. Sali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology* 234(3):779-815.

14. Song Y, *et al.* (2013) High-resolution comparative modeling with RosettaCM. *Structure* 21(10):1735-1742.

15. Groll M, Brandstetter H, Bartunik H, Bourenkow G, & Huber R (2003) Investigations on the maturation and regulation of archaebacterial proteasomes. *Journal of molecular biology* 327(1):75-83.

16. Trabuco LG, Villa E, Schreiner E, Harrison CB, & Schulten K (2009) Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* 49(2):174-180.

17. Ribeiro JV, *et al.* (2016) QwikMD - Integrative Molecular Dynamics Toolkit for Novices and Experts. *Scientific reports* 6:26536.

18. Humphrey W, Dalke A, & Schulten K (1996) VMD: visual molecular dynamics. *Journal of molecular graphics* 14(1):33-38, 27-38.

19. Phillips JC, *et al.* (2005) Scalable molecular dynamics with NAMD. *Journal of computational chemistry* 26(16):1781-1802.

20. Lindert S & McCammon JA (2015) Improved cryoEM-Guided Iterative Molecular Dynamics--Rosetta Protein Structure Refinement Protocol for High Precision Protein Structure Prediction. *Journal of chemical theory and computation* 11(3):1337-1346.

21. Pettersen EF, *et al.* (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* 25(13):1605-1612.